

# Compressed Sensing with Nonlinear Observations and Related Nonlinear Optimisation Problems

Thomas Blumensath

University of Oxford

Centre for Functional Magnetic Resonance Imaging of the Brain

J R Hospital, Oxford, OX3 9DU, UK

tblumens@fmrib.ox.ac.uk

## Abstract

Non-convex constraints have recently proven a valuable tool in many optimisation problems. In particular sparsity constraints have had a significant impact on sampling theory, where they are used in Compressed Sensing and allow structured signals to be sampled far below the rate traditionally prescribed.

Nearly all of the theory developed for Compressed Sensing signal recovery assumes that samples are taken using linear measurements. In this paper we instead address the Compressed Sensing recovery problem in a setting where the observations are non-linear. We show that, under conditions similar to those required in the linear setting, the Iterative Hard Thresholding algorithm can be used to accurately recover sparse or structured signals from few non-linear observations.

Similar ideas can also be developed in a more general non-linear optimisation framework. In the second part of this paper we therefore present related result that show how this can be done under sparsity and union of subspaces constraints, whenever a generalisation of the Restricted Isometry Property traditionally imposed on the Compressed Sensing system holds.

*Key words and phrases* : Compressed Sensing, Nonlinear Optimisation, Non-Convex Constraints, Inverse Problems

## 1 Introduction

Compressed Sensing [1, 2, 3] deals with the acquisition of finite dimensional sparse signals. Let  $\mathbf{x}$  be a sparse vector of length  $N$  and assume we sample  $\mathbf{x}$  using  $M$  *linear* measurements. The  $M$  samples can then be collected into a vector  $\mathbf{y}$  of length  $M$  and the sampling process can be described by a matrix  $\Phi$ .

If the observations are noisy, then the Compressed Sensing observation model is

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}, \quad (1)$$

where  $\mathbf{e}$  is the noise vector. If  $M < N$ , then such a linear system is not uniquely invertible in general, unless we use additional assumptions on  $\mathbf{x}$ . Sparsity of  $\mathbf{x}$  is such an assumption and Compressed Sensing theory tells us that, for certain  $\Phi$ , we can recover  $\mathbf{x}$  from  $\mathbf{y}$  even if  $M \ll N$ , given that  $\mathbf{x}$  has roughly  $O(M)$  non-zero elements. However, in general, recovery of  $\mathbf{x}$  is a combinatorial problem which is known to be NP-hard. Fortunately, under stricter conditions on  $\Phi$ , a range of different polynomial time algorithms can be used to recover  $\mathbf{x}$  whenever  $\mathbf{x}$  has roughly  $O(M/\log(N))$  non-zero elements.

One of the conditions that guarantees that we can use efficient algorithms is the *Restricted Isometry Property* (RIP). A matrix  $\Phi$  satisfies the *Restricted Isometry Property* of order  $2k$  [1] if

$$(1 - \delta)\|\mathbf{x}_1 + \mathbf{x}_2\|^2 \leq \|\Phi(\mathbf{x}_1 + \mathbf{x}_2)\|^2 \leq (1 + \delta)\|\mathbf{x}_1 + \mathbf{x}_2\|^2 \quad (2)$$

for all  $k$ -sparse  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . The *Restricted Isometry Constant*  $\delta$  is defined as the smallest constant for which this property holds. One important interpretation of the RIP is in terms of the Lipschitz property of  $\Phi$  and its inverse (where the inverse is defined only for sparse vectors and their image under  $\Phi$ ) [14] and the condition states that, not only is  $\Phi$  invertible on the set of sparse signals, this inverse is also smooth.

The RIP condition is a sufficient condition for the recovery of sparse  $\mathbf{x}$ . For example, [4] has shown that, for any  $\mathbf{x}$ , given an observation  $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$ , where  $\Phi$  has the Restricted Isometry Property with  $\delta < \sqrt{2} - 1$ , then the solution  $\mathbf{x}^*$  to the convex optimisation problem

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1 : \|\mathbf{y} - \Phi \tilde{\mathbf{x}}\|_2 \leq \|\mathbf{e}\|_2 \quad (3)$$

has an error bounded by

$$\|\mathbf{x}^* - \mathbf{x}\| \leq ck^{-0.5}\|\mathbf{x} - \mathbf{x}_k\|_1 + c'\|\mathbf{e}\|, \quad (4)$$

where  $\|\cdot\|_1$  is the vector 1 norm,  $\mathbf{x}_k$  is the best  $k$  term approximation to  $\mathbf{x}$  and where  $c$  and  $c'$  are two constants depending only on  $\delta$ .

Similar results have been obtained for other algorithms, such as the Compressed Sampling Matching Pursuit (CoSaMP) and Subspace Pursuit (SP) algorithms [5, 6] and the Iterative Hard Thresholding (IHT) algorithm [7].

Several generalisations to this now classical Compressed Sensing setup have been introduced over the years. For example, some of the recovery algorithms, such as CoSaMP, SP and IHT, can be adapted to allow signals  $\mathbf{x}$  to lie in a much more general, *non-convex* constraint set  $\mathcal{A}$ . A powerful model here is for

example the Union of Subspaces model, in which  $\mathbf{x}$  is assumed to lie on one of several linear subspaces  $\mathcal{A}_i$ , though it is not known a priori on which subspace we are to look. Not only does this framework include the standard sparse model as a special instance, many other models of interest, such as analogue Compressed Sensing methods [8], low rank matrix models [9], or structured sparse models [10], are also covered.

In this more general setting, with a general *non-convex* constraint sets  $\mathcal{A}$ , Compressed Sensing can be formulated as the following optimisation problem,

$$\operatorname{argmin}_{\mathbf{x} \in \mathcal{A}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2, \quad (5)$$

that is, we search a vector  $\mathbf{x}$  from the *non-convex* constraint set  $\mathcal{A}$  that minimises the sum of squares observation error.

In this paper we look at a much more general setting, where we want to find the following optimum.

$$\operatorname{argmin}_{\mathbf{x} \in \mathcal{A}} f(\mathbf{x}), \quad (6)$$

where  $f(\mathbf{x})$  is now a much more general non-linear function of  $\mathbf{x}$ .

Of particular interest to us are *non-linear Compressed Sensing* problems where  $f(\mathbf{x}) = \|\mathbf{y} - \Phi(\mathbf{x})\|$ , with  $\Phi(\mathbf{x})$  being a non-linear mapping from one vector space to another. We address this non-linear Compressed Sensing problem first, however, the more general problem in equation (6) is of independent interest and an alternative treatment will be presented in the second part of this paper.

When we started studying these problems, not much was known of this general setting. However, since the first draft of this paper [11], similar ideas have been put forward independently in [12], where the non-linear Compressed Sensing problem was tackled using a convexification approach, and in [13], where non-convex optimisation problems were studied using an alternative greedy approach to the one discussed here. Whilst the first part of this paper contains more recent results, the second part of this paper is basically the same material that can be found in the earlier draft of this paper [11].

## 2 Non-Linear Compressed Sensing

We are here interested in the development of a better understanding of what happens to the Compressed Sensing recovery problem when a signal is measured with some non-linear system. In particular, the hope is that, if the system is not too non-linear, then recovery should still be possible under similar assumption to those made in linear Compressed Sensing. To see the intuition behind why this might work, it is worth pointing out that in the linear setting, Compressed Sensing recovery works exactly in those cases in which the observation system is a bi-Lipschitz embedding. This means that, both, the observation mapping itself, as well as its inverse are Lipschitz functions. Obviously, these functions

are only Lipschitz on the constraint set  $\mathcal{A}$  and its image  $\Phi\mathcal{A}$ . In the linear setting, if  $\Phi$  is bounded (e.g. in finite dimensional spaces), then  $\Phi$  itself is obviously Lipschitz. The idea is now that, if Compressed Sensing works if both forward and backward maps are Lipschitz, maybe we can move away from a linear setting, where  $\Phi$  is linear, and instead assume  $\Phi$  to be Lipschitz, but non-linear.

The study of non-linear observation systems is not only of academic interest but has important implications for many real-world sampling systems, where measurement system can often not be designed to be perfectly linear. Assume therefore that our measurements are described by a nonlinear mapping  $\Phi(\cdot)$  that maps elements of the normed vector spaces  $\mathcal{H}$  into the normed vector spaces  $\mathcal{B}$ . The observation model is therefore

$$\mathbf{y} = \Phi(\mathbf{x}) + \mathbf{e}, \quad (7)$$

where  $\mathbf{e} \in \mathcal{B}$  is an unknown but bounded error term. Both  $\mathcal{H}$  and  $\mathcal{B}$  are assumed to be Hilbert spaces.

## 2.1 The Constraints

As in Compressed Sensing, the interesting case occurs whenever the sampling system  $\Phi$  is non-invertible or ill-conditioned. To cope with this, additional constraints need to be imposed on  $\mathbf{x}$ . Again, in the interest of generality, instead of restricting our discussion to sparse signals (however these might be defined in a general Hilbert spaces) we here use the more general framework of [14] and assume that  $\mathbf{x}$  lies in or close to a known set  $\mathcal{A}$ , where  $\mathcal{A} \subset \mathcal{H}$  is a non-convex subset of  $\mathcal{H}$ . Of particular interest will be constraint sets  $\mathcal{A}$  that can be described as the union of several subspaces. For these models we can write

$$\mathcal{A} = \bigcup \mathcal{A}_i, \quad (8)$$

where we use arbitrary closed subspaces  $\mathcal{A}_i \subset \mathcal{H}$

One approach to recover  $\mathbf{x}$  from  $\mathbf{y}$  would be to mirror Compressed Sensing ideas and to define a convex objective function which can then be optimised using standard tools. However, for our general setup, it is not clear how this could be done. Instead, we use the Iterative Hard Thresholding (IHT) algorithm. To define this for general constraint sets  $\mathcal{A}$ , we again replace the hard thresholding step with a more general map which can be understood as a form of projection [14]. Let  $P_{\mathcal{A}}$  be a map from  $\mathcal{H}$  to  $\mathcal{A}$  such that

$$\mathbf{x}_{\mathcal{A}} = P_{\mathcal{A}}(\mathbf{x}) : \mathbf{x}_{\mathcal{A}} \in \mathcal{A}, \|\mathbf{x} - \mathbf{x}_{\mathcal{A}}\|^2 \leq \inf_{\hat{\mathbf{x}} \in \mathcal{A}} \|\mathbf{x} - \hat{\mathbf{x}}\|^2 + \epsilon. \quad (9)$$

In this definition we have introduced an arbitrarily small constant  $\epsilon > 0$ , as there might not exist an  $\mathbf{x}_{opt}$ , such that  $\|\mathbf{x} - \mathbf{x}_{opt}\|^2 = \inf_{\hat{\mathbf{x}} \in \mathcal{A}} \|\mathbf{x} - \hat{\mathbf{x}}\|^2$ . However, for

simplicity, we will assume for the rest of this paper that  $\mathcal{A}$  is a so called proximal set, which is just a fancy way of saying that the required optimal points indeed lie in the set  $\mathcal{A}$ , so that we use  $\epsilon = 0$  here. Nevertheless, it is easy to adapt our theory to the more general setting.

Note that this "projection" might not be defined uniquely in general, as for a given  $\mathbf{x}$ , there might be several elements  $\mathbf{x}_{\mathcal{A}}$  that satisfy the condition in (9). However, all we require here is that the map  $P_{\mathcal{A}}(\mathbf{x})$  returns a *single element* from the set of admissible  $\mathbf{x}_{\mathcal{A}}$  (which is guaranteed to be non-empty [14]). How this selection is done is of no consequence for our arguments here.

It is further worth noting that the relaxation offered by an  $\epsilon > 0$  in the definition of the above projection has also a computational advantage. Instead of having to compute exact optima, which for many problems are often difficult to find, many approximate algorithms can be used instead (see [15] for a more detailed discussion).

## 2.2 The Iterative Hard Thresholding Algorithm for Non-Linear Compressed Sensing

For the linear Compressed Sensing problem, the Iterative Hard Thresholding (IHT) algorithm uses the following iteration

$$\mathbf{x}^{n+1} = P_{\mathcal{A}}(\mathbf{x}^n + \mu \Phi^*(\mathbf{y} - \Phi \mathbf{x}^n)), \quad (10)$$

where  $\Phi$  is the linear measurement operator.

In the non-linear case, let us approximate  $\Phi(\mathbf{x})$  using an affine Taylor series type approximation around a point  $\mathbf{x}^*$ , so that  $\Phi(\mathbf{x}) \approx \Phi(\mathbf{x}^*) + \Phi_{\mathbf{x}^*}(\mathbf{x} - \mathbf{x}^*)$ , where  $\Phi_{\mathbf{x}^*}$  is a linear operator (such as the Jacobian of  $\Phi(\mathbf{x})$ , evaluated at point  $\mathbf{x}^*$ ). The matrix  $\Phi_{\mathbf{x}^*}$  thus will depend on  $\mathbf{x}^*$  in general. At iteration  $n$  we then write the IHT algorithm as

$$\mathbf{x}^{n+1} = P_{\mathcal{A}}(\mathbf{x}^n + \mu \Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi(\mathbf{x}^n))). \quad (11)$$

Indeed, as we show below in 2.4, this algorithm can recover  $\mathbf{x}$  under similar condition to those required from the IHT algorithm in the linear setting. All we require is that the matrices  $\Phi_{\mathbf{x}^*}$  satisfy a Restricted Isometry Property and that the error introduced in the linearisation is not too large, i.e. that  $\|\Phi(\mathbf{x}_{\mathcal{A}}) - \Phi(\mathbf{x}_{\mathcal{A}}^n) - \Phi_{\mathbf{x}_{\mathcal{A}}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}_{\mathcal{A}}^n)\|$  is small for large  $n$ .

**Theorem 1.** *Assume that  $\mathbf{y} = \Phi(\mathbf{x}) + \mathbf{e}$  and that  $\Phi_{\mathbf{x}^*}$  is a linearisation of  $\Phi(\cdot)$  at  $\mathbf{x}^*$  so that the Iterative Hard Thresholding algorithm uses the iteration  $\mathbf{x}^{n+1} = P_{\mathcal{A}}(\mathbf{x}^n + \mu \Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi_{\mathbf{x}^n} \mathbf{x}^n))$ . Assume that  $\Phi_{\mathbf{x}^*}$  satisfies RIP*

$$\alpha \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \leq \alpha \|\Phi_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq \beta \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \quad (12)$$

for all  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}^* \in \mathcal{A}$ , with constants satisfying  $\beta \leq 1/\mu < 1.5\alpha$ . Define

$$\mathbf{e}_{\mathcal{A}}^n = \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n) \quad (13)$$

and  $\epsilon^k = b \sum_{n=0}^{k-1} a^{k-1-n} \|\mathbf{e}_{\mathcal{A}}^n\|^2$ , where  $b = 4/\alpha$  and  $a = 2/(\mu\alpha) - 2$ , then after

$$k^* = \left\lceil 2 \frac{\ln(\delta \frac{\sqrt{\epsilon^k}}{\|\mathbf{x}_{\mathcal{A}}\|})}{\ln(2/(\mu\alpha) - 2)} \right\rceil \quad (14)$$

iterations we have

$$\|\mathbf{x} - \mathbf{x}^{k^*}\| \leq (1 + \delta)\sqrt{\epsilon^k} + \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}\|. \quad (15)$$

Obviously, for the above theorem to make sense, we would require the error term  $\epsilon^k$  to be well behaved. This is true whenever  $\|\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|_2$  is bounded, as then  $\epsilon^k \leq b \sum_{n=0}^{k-1} a^{k-1-n} C$ , for some constant  $C$  so that the requirement that  $1/\mu < 1.5\alpha$  ensures that  $a = 2/(\mu\alpha) - 2 < 1$ , which in turn implies that the geometric series  $\sum_{n=0}^{k-1} a^{k-1-n}$  is bounded.

Indeed, if  $a < 1$  and if we can show that  $\epsilon_n = \|\mathbf{e}_{\mathcal{A}}^n\|^2$  is bounded and convergent to some  $\epsilon_{lim}$ , then  $\epsilon^k$  will also be bounded as the following argument shows

$$\begin{aligned} \epsilon^k/b &= \sum_{n=0}^{k-1} a^{k-n-1} \epsilon_n \\ &= \sum_{n=0}^{p-1} a^{k-n-1} \epsilon_n + \sum_p^{k-1} a^{k-n-1} \epsilon_n \\ &\leq \sum_{n=0}^{p-1} a^{k-n-1} \epsilon_n + \epsilon_p \sum_p^{k-1} a^{k-n-1} \\ &\leq \sum_{n=0}^{p-1} a^{k-n-1} \epsilon_n + \epsilon_p \frac{1}{1-a} \\ &= a^{k-p-1} \sum_{n=0}^{p-1} a^{p-n-1} \epsilon_n + \epsilon_p \frac{1}{1-a} \\ &\leq a^{k-p-1} \epsilon_0 \sum_{n=0}^{p-1} a^{p-n-1} + \epsilon_p \frac{1}{1-a} \\ &\leq a^{k-p-1} \frac{\epsilon_0}{1-a} + \frac{\epsilon_p}{1-a} \end{aligned} \quad (16)$$

Thus, if we let  $k$  and  $p$  increase to infinity such that  $0 < k - p \rightarrow \infty$ , then the first term on the left converges to zero whilst the second term converges to a

limit depending on  $\epsilon_{lim}$ , so that, if we iterate the algorithm long enough, then

$$\lim_{k \rightarrow \infty} \epsilon^k \leq \epsilon_{lim} \frac{b}{1-a} \quad (17)$$

and the error term converges to

$$\|\mathbf{x} - \mathbf{x}^*\| \leq \sqrt{\epsilon_{lim} \frac{b}{1-a}} + \|\mathbf{x}_A - \mathbf{x}\|. \quad (18)$$

Actually, as shown in 2.5, more can be said if we can establish the following bound for  $\Phi(\mathbf{x})$  and its linearisation  $\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq C\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$ .

**Corollary 2.** *Assume that  $\mathbf{y} = \Phi(\mathbf{x}) + \mathbf{e}$  and that  $\Phi_{\mathbf{x}^*}$  is a linearisation of  $\Phi(\cdot)$  at  $\mathbf{x}^*$  so that the Iterative Hard Thresholding algorithm uses the iteration  $\mathbf{x}^{n+1} = P_A(\mathbf{x}^n + \mu\Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi_{\mathbf{x}^n}\mathbf{x}^n)$ . Assume that  $\Phi_{\mathbf{x}^*}$  satisfies RIP*

$$\alpha\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \leq \|\Phi_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq \beta\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \quad (19)$$

for all  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}^* \in \mathcal{A}$ , and assume  $\Phi(\mathbf{x})$  and  $\Phi_{\mathbf{x}}$  satisfy

$$\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq C\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2, \quad (20)$$

with constants satisfying  $\beta \leq 1/\mu < 1.5\alpha - 4C$ , then the algorithm converges to a solution  $\mathbf{x}^*$  that satisfies

$$\|\mathbf{x} - \mathbf{x}^*\| \leq c\|\mathbf{e}_A\| + \|\mathbf{x}_A - \mathbf{x}\|, \quad (21)$$

where  $\mathbf{e}_A = \mathbf{y} - \Phi(\mathbf{x}_A)$  and  $c = \frac{2}{0.75\alpha - 1/\mu - 2C}$ .

### 2.3 Example

Before we proof Theorem 1 and Corollary 2, let us give a simple example that shows how the above method and theory can be applied in a particular setting. Assume we have constructed a Compressed Sensing system, where a sparse signal  $\mathbf{x} \in \mathbb{R}^N$  is measured using a *linear* measurement system  $\bar{\Phi}$ . Also assume that we have constructed the system so that  $\bar{\Phi}$  satisfies the Restricted Isometry Property with constants  $\alpha$  and  $\beta$ . Now unfortunately, the sensors we have available for the actual measurements are not exactly linear but have a slight non-linearity, so that our measurements are of the form

$$\mathbf{y} = \bar{\Phi}f(\mathbf{x}) + \mathbf{e}, \quad (22)$$

where  $f(\cdot)$  is a non-linear function applied to each element of the vector  $\mathbf{x}$ . For simplicity, we will write  $f(\mathbf{x}) = \mathbf{x} + h(\mathbf{x})$ , where again  $h(\mathbf{x})$  is a function applied element wise. We then have  $f'(x) = 1 + h'(x)$ .

It is not difficult to see that the Jacobian of  $\overline{\Phi}f(\mathbf{x})$  can be written as

$$\Phi_{\mathbf{x}^*} = \overline{\Phi} + \overline{\Phi}H'_{\mathbf{x}^*}, \quad (23)$$

where  $H'_{\mathbf{x}^*}$  is the diagonal matrix with the elements  $h'(\mathbf{x}^*)$  along the diagonal.

To use Corollary 2, we thus need to determine a) the RIP constant of  $\Phi_{\mathbf{x}^*}$  and b) bound  $\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2$  as a function of  $\|\mathbf{x}_1 - \mathbf{x}_2\|^2$ .

The RIP constants are bounded for our example as follows. Assume that  $|h'(x)| \leq M < 1$ , that  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{A}$  and that  $\overline{\Phi}$  satisfies the RIP with constants  $\alpha$  and  $\beta$ . We then have

$$\begin{aligned} & (\alpha^{1/2}1 - \beta^{1/2}M)\|\mathbf{x}_1 - \mathbf{x}_2\| \\ & \leq \|\overline{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\| - \|\overline{\Phi}H'_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\| \\ & \leq \|\Phi_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\| \\ & = \|\overline{\Phi}(\mathbf{x}_1 - \mathbf{x}_2) + \overline{\Phi}H'_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\| \\ & \leq \|\overline{\Phi}(\mathbf{x}_1 - \mathbf{x}_2)\| + \|\overline{\Phi}H'_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\| \\ & \leq \beta\|\mathbf{x}_1 - \mathbf{x}_2\| + \beta\|H'_{\mathbf{x}^*}(\mathbf{x}_1 - \mathbf{x}_2)\| \\ & \leq \beta^{1/2}(1 + M)\|\mathbf{x}_1 - \mathbf{x}_2\|, \end{aligned} \quad (24)$$

which proofs the following Lemma.

**Lemma 3.** *Let  $\Phi(\mathbf{x}) = \overline{\Phi}f(\mathbf{x})$ , where the function  $f(\mathbf{x}) = x + h(\mathbf{x})$  is applied element wise and where the derivative  $h'(x)$  is absolutely bounded  $|h'(\cdot)| \leq M$ . Also assume that the matrix  $\overline{\Phi}$  satisfies the RIP condition with constants  $\alpha$  and  $\beta$  for a set  $\mathcal{A}$ , then the matrix  $\overline{\Phi}(\mathbf{I} + H'_{\mathbf{x}^*})$  satisfies RIP with constants  $(\alpha^{1/2}1 - \beta^{1/2}M)^2$  and  $\beta(1 - M)^2$ .*

Let us now turn to point b). We have the bound

$$\begin{aligned} & \|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \\ & = \|\overline{\Phi}h(\mathbf{x}_1) - \overline{\Phi}h(\mathbf{x}_2) - \overline{\Phi}H'_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \\ & = \|\overline{\Phi}(h(\mathbf{x}_1) - h(\mathbf{x}_2) - H'_{\mathbf{x}_1}\mathbf{x}_1 + H'_{\mathbf{x}_1}\mathbf{x}_2)\|_2^2 \\ & \leq \beta\|(h(\mathbf{x}_1) - h(\mathbf{x}_2) - H'_{\mathbf{x}_1}\mathbf{x}_1 + H'_{\mathbf{x}_1}\mathbf{x}_2)\|_2^2, \end{aligned} \quad (25)$$

where in the last inequality we assume  $\overline{\Phi}$  to satisfy the RIP property and that  $h(0) = 0$  (Note that if we do not assume that  $h(0) = 0$ , then the same results still hold, though we have to replace  $\beta$  by the operator norm of  $\overline{\Phi}$ ). Let us introduce the function  $d_{\mathbf{x}^*}(\mathbf{x}) = h(\mathbf{x}) - H'_{\mathbf{x}^*}\mathbf{x}$ , so that

$$\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq \beta\|d_{\mathbf{x}_1}(\mathbf{x}_1) - d_{\mathbf{x}_1}(\mathbf{x}_2)\|_2^2. \quad (26)$$

Thus if  $d_{\mathbf{x}^*}$  is Lipschitz for all  $\mathbf{x}^* \in \mathcal{A}$  with a small constant  $K$ , then the condition

$$\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq C\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2, \quad (27)$$



in Corollary 2 holds with  $C = \beta K$ .

Thus it remains to show that  $d_{\mathbf{x}_1}(\mathbf{x}_1)$  is Lipschitz. If the Jacobian  $D_{\mathbf{x}_1}(\mathbf{x}^*)$  of  $d_{\mathbf{x}_1}(\mathbf{x}_1)$  satisfies  $\|D_{\mathbf{x}_1}(\mathbf{x} + t\mathbf{h})\| \leq M$  for all  $0 \leq t \leq 1$ , then we know that

$$\|d_{\mathbf{x}_1}(\mathbf{x} + \mathbf{h}) - d_{\mathbf{x}_1}(\mathbf{x})\| \leq M\|\mathbf{h}\|, \quad (28)$$

so that

$$\|d_{\mathbf{x}_1}(\mathbf{x}_1) - d_{\mathbf{x}_1}(\mathbf{x}_2)\| = \|d_{\mathbf{x}_1}(\mathbf{x}_2 + \mathbf{x}_1 - \mathbf{x}_2) - d_{\mathbf{x}_1}(\mathbf{x}_2)\| \leq M\|\mathbf{x}_1 - \mathbf{x}_2\|, \quad (29)$$

holds if

$$\|D_{\mathbf{x}_1}(\mathbf{x}_2 + t(\mathbf{x}_1 - \mathbf{x}_2))\| \leq M \quad (30)$$

for all  $0 \leq t \leq 1$ .

For our simple example, we see that  $D_{\mathbf{x}_1}$  is in fact a diagonal matrix with entries  $\{D_{\mathbf{x}_1}(\mathbf{x}^*)\}_{i,i} = h'(x_i) - h'_i(\mathbf{x}^*)$ , where  $h'_i(\mathbf{x}_1) = h'(\{\mathbf{x}_1\}_i)$ , so that

$$\{D_{\mathbf{x}_1}(\mathbf{x}_2 + t(\mathbf{x}_1 - \mathbf{x}_2))\}_{i,i} = h'(\{\mathbf{x}_2 + t(\mathbf{x}_1 - \mathbf{x}_2)\}_i) - h'(\{\mathbf{x}_1\}_i). \quad (31)$$

Thus if  $|h'(\cdot)|$  is bounded, that is, if  $|h'(\cdot)| \leq M$ , then  $\|D_{\mathbf{x}_1}(\mathbf{x})\| \leq M$ .

We thus have demonstrated the following.

**Lemma 4.** *Let  $\Phi(\mathbf{x}) = \overline{\Phi}f(\mathbf{x})$ , where the function  $f(\cdot) = x + h(x)$  is applied element wise and where the derivative  $h'(x)$  is absolutely bounded  $|h'(\cdot)| \leq M$ , then*

$$\|\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2) - \Phi_{\mathbf{x}_1}(\mathbf{x}_1 - \mathbf{x}_2)\|_2^2 \leq C\|\mathbf{x}_1 - \mathbf{x}_2\|, \quad (32)$$

where  $C = \beta M$ .

## 2.4 Proof of Theorem 1

The proof follows basically that in [14], but with some important modifications to account for the non-linear setting analysed here.

*Proof.* As always, we start with the triangle inequality

$$\|\mathbf{x} - \mathbf{x}^{n+1}\| \leq \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1}\| + \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}\| \quad (33)$$

and then bound the first term on the left using the definition

$$\mathbf{e}_{\mathcal{A}}^n = \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n) \quad (34)$$

and the inequalities

$$\begin{aligned}
& \| \mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1} \|^2 \\
\leq & \frac{1}{\alpha} \| \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1}) \|^2 \\
= & \frac{1}{\alpha} \| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) - (\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)) \|^2 \\
= & \frac{1}{\alpha} \| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) - \mathbf{e}_{\mathcal{A}}^n \|^2 \\
= & \frac{1}{\alpha} (\| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \|^2 + \| \mathbf{e}_{\mathcal{A}}^n \|^2 - 2 \langle \mathbf{e}_{\mathcal{A}}^n, (\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n)) \rangle) \\
\leq & \frac{1}{\alpha} (\| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \|^2 + \| \mathbf{e}_{\mathcal{A}}^n \|^2 + \| \mathbf{e}_{\mathcal{A}}^n \|^2 + \| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \|^2) \\
= & \frac{2}{\alpha} \| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \|^2 + \frac{2}{\alpha} \| \mathbf{e}_{\mathcal{A}}^n \|^2. \tag{35}
\end{aligned}$$

We here used the fact that

$$\begin{aligned}
& - \langle \mathbf{e}_{\mathcal{A}}^n, (\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n)) \rangle \\
\leq & \| \mathbf{e}_{\mathcal{A}}^n \| \| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \| \\
\leq & 0.5 (\| \mathbf{e}_{\mathcal{A}}^n \|^2 + \| (\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n)) \|^2).
\end{aligned}$$

The left term in the last line of (35) is bounded by the next inequality

$$\| \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n) \|^2 \leq \left( \frac{1}{\mu} - \alpha \right) \| (\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n) \|^2 + \| \mathbf{e}_{\mathcal{A}}^n \|^2, \tag{36}$$

which is a result of the following argument in which we use  $\mathbf{g} = 2\Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi(\mathbf{x}^n))$

$$\begin{aligned}
& \|\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n)\|^2 - \|\mathbf{y} - \Phi(\mathbf{x}^n)\|^2 \\
&= -\langle(\mathbf{x}^{n+1} - \mathbf{x}^n), \mathbf{g}\rangle + \|\Phi_{\mathbf{x}^n}(\mathbf{x}^{n+1} - \mathbf{x}^n)\|^2 \\
&\leq -\frac{2}{\mu}\langle(\mathbf{x}^{n+1} - \mathbf{x}^n), \frac{\mu}{2}\mathbf{g}\rangle + \frac{1}{\mu}\|(\mathbf{x}^{n+1} - \mathbf{x}^n)\|^2 \\
&= \frac{1}{\mu}\left[\|\mathbf{x}^{n+1} - \mathbf{x}^n - \frac{\mu}{2}\mathbf{g}\|^2 - \frac{\mu}{2}\|\mathbf{g}\|^2\right] \\
&= \frac{1}{\mu}\left[\inf_{\mathbf{x} \in \mathcal{A}} \|\mathbf{x} - \mathbf{x}^n - \frac{\mu}{2}\mathbf{g}\|^2 - \frac{\mu}{2}\|\mathbf{g}\|^2\right] \\
&= \inf_{\mathbf{x} \in \mathcal{A}} \left[-\langle(\mathbf{x} - \mathbf{x}^n), \mathbf{g}\rangle + \frac{1}{\mu}\|(\mathbf{x} - \mathbf{x}^n)\|^2\right] \\
&\leq -\langle(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n), \mathbf{g}\rangle + \frac{1}{\mu}\|(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 \\
&= -2\langle(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n), \Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi(\mathbf{x}^n))\rangle + \frac{1}{\mu}\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2 \\
&= -2\langle(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n), \Phi_{\mathbf{x}^n}^*(\mathbf{y} - \Phi(\mathbf{x}^n))\rangle + \alpha\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2 \\
&\quad + \left(\frac{1}{\mu} - \alpha\right)\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2 \\
&\leq -2\langle(\Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n), (\mathbf{y} - \Phi(\mathbf{x}^n)))\rangle + \|\Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 \\
&\quad + \left(\frac{1}{\mu} - \alpha\right)\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2 \\
&= \|\mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 - \|\mathbf{y} - \Phi(\mathbf{x}^n)\|^2 \\
&\quad + \left(\frac{1}{\mu} - \alpha\right)\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2 \\
&= \|\mathbf{e}_{\mathcal{A}}^n\|^2 - \|\mathbf{y} - \Phi(\mathbf{x}^n)\|^2 + \left(\frac{1}{\mu} - \alpha\right)\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2, \tag{37}
\end{aligned}$$

where the first and last inequalities are due to the RIP property of  $\Phi_{\mathbf{x}^n}$  and the choice of  $\beta \leq \frac{1}{\mu}$ , whilst the second inequality is due to the fact that  $\mathbf{x}_{\mathcal{A}} \in \mathcal{A}$ .

We have thus shown that

$$\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1}\|^2 \leq 2\left(\frac{1}{\mu\alpha} - 1\right)\|(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 + \frac{4}{\alpha}\|\mathbf{e}_{\mathcal{A}}^n\|^2. \tag{38}$$

We can now iterate the above expression. Using  $\epsilon^k = b \sum_{n=0}^{k-1} \|\mathbf{y} - \Phi_{\mathbf{x}^n} \mathbf{x}_{\mathcal{A}}\|^2 a^{k-1-n}$ , where  $b = 4/\alpha$  and  $a = 2/(\mu\alpha) - 2$ , we get

$$\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^k\|^2 \leq \left(2\left(\frac{1}{\mu\alpha} - 1\right)\right)^k \|\mathbf{x}_{\mathcal{A}}\|^2 + \epsilon^k. \tag{39}$$

Thus

$$\begin{aligned}\|\mathbf{x} - \mathbf{x}^k\| &\leq \sqrt{c^k \|\mathbf{x}_{\mathcal{A}}\|^2 + \epsilon^k + \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}\|} \\ &\leq c^{k/2} \|\mathbf{x}_{\mathcal{A}}\| + \sqrt{\epsilon^k} + \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}\|,\end{aligned}$$

where  $c = \frac{2}{\mu\alpha} - 2$  and the theorem is proven.  $\square$

## 2.5 Proof of Corollary 2

*Proof.* Let us start with the bound in (38)

$$\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1}\|^2 \leq 2 \left( \frac{1}{\mu\alpha} - 1 \right) \|(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 + \frac{4}{\alpha} \|\mathbf{e}_{\mathcal{A}}^n\|^2 \quad (40)$$

and let us look a bit more closely at

$$\begin{aligned}\mathbf{e}_{\mathcal{A}}^n &= \mathbf{y} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n) \\ &= \Phi(\mathbf{x}_{\mathcal{A}}) + \mathbf{e}_{\mathcal{A}} - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n),\end{aligned} \quad (41)$$

where  $\mathbf{e}_{\mathcal{A}} = \mathbf{y} - \Phi(\mathbf{x}_{\mathcal{A}})$ . We then have

$$\|\mathbf{e}_{\mathcal{A}}^n\| \leq \|\Phi(\mathbf{x}_{\mathcal{A}}) - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\| + \|\mathbf{e}_{\mathcal{A}}\| \quad (42)$$

Now by assumption,  $\|\Phi(\mathbf{x}_{\mathcal{A}}) - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|$  is bounded as a function of  $\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|$ , i.e.

$$\|\Phi(\mathbf{x}_{\mathcal{A}}) - \Phi(\mathbf{x}^n) - \Phi_{\mathbf{x}^n}(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 \leq C \|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n\|^2, \quad (43)$$

so that (38) becomes

$$\|\mathbf{x}_{\mathcal{A}} - \mathbf{x}^{n+1}\|^2 \leq 2 \left( \frac{1}{\mu\alpha} - 1 + \frac{4}{\alpha} C \right) \|(\mathbf{x}_{\mathcal{A}} - \mathbf{x}^n)\|^2 + \frac{8}{\alpha} \|\mathbf{e}_{\mathcal{A}}\|^2, \quad (44)$$

Thus we require that  $\frac{1}{\mu\alpha} - 1 + \frac{4}{\alpha} C \leq 0.5$ , that is that  $1/\mu \leq 1.5\alpha - 4C$ .

The same argument used in the main proof now holds. Whenever the constant before the left term on the right hand side is smaller than one, then we can iterate the error and the corollary follows.  $\square$

## 3 The Iterative Hard Thresholding Algorithm for Non-Linear Optimisation

Let us now return to the more general problem of minimising a non-linear function  $f(\mathbf{x})$  under the constraint that  $\mathbf{x} \in \mathcal{A}$ , where  $\mathcal{A}$  is a Union of Subspaces. Let

us recall again that for minimisation problems of the form  $\operatorname{argmin}_{\mathbf{x} \in \mathcal{A}} \|\mathbf{y} - \Phi \mathbf{x}\|^2$  we use the algorithm

$$\mathbf{x}^{n+1} = P_{\mathcal{A}}(\mathbf{x}^n + \Phi^*(\mathbf{y} - \Phi \mathbf{x}^n)). \quad (45)$$

Note that the update  $\Phi^*(\mathbf{y} - \Phi \mathbf{x}^n)$  is a scaled version of the gradient of the cost function  $\|\mathbf{y} - \Phi \mathbf{x}\|^2$ .

In the more general setting  $\operatorname{argmin}_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ , where  $\mathbf{x}$  is an Euclidean vector, we can simply replace this update direction with the gradient of  $f(\mathbf{x})$  (evaluated at  $\mathbf{x}^n$ ), whilst in more general spaces, we assume that  $f(\mathbf{x})$  is Fréchet differentiable with respect to  $\mathbf{x}$ , that is, for each  $\mathbf{x}_1$  there exist a linear functional  $D_{\mathbf{x}_1}(\cdot)$  such that

$$\lim_{\mathbf{h} \rightarrow 0} \frac{f(\mathbf{x}_1 + \mathbf{h}) - f(\mathbf{x}_1) - D_{\mathbf{x}_1}(\mathbf{h})}{\|\mathbf{h}\|} = 0. \quad (46)$$

We can then use Riesz representation theorem to write the linear functional  $D_{\mathbf{x}_1}(\cdot)$  using its inner product equivalent

$$D_{\mathbf{x}_1}(\cdot) = \langle \nabla(\mathbf{x}_1), \cdot \rangle, \quad (47)$$

where  $\nabla(\mathbf{x}_1) \in \mathcal{H}$ . Using  $\mathbf{h} = \mathbf{x}_2 - \mathbf{x}_1$  we see that for each  $\mathbf{u}$  and  $\mathbf{x}_1$  we require the existence of a  $\nabla(\mathbf{x}_1)$  such that

$$\lim_{\mathbf{x}_2 \rightarrow \mathbf{x}_1} \frac{f(\mathbf{x}_2) - f(\mathbf{x}_1) - \langle \nabla(\mathbf{x}_1), (\mathbf{x}_2 - \mathbf{x}_1) \rangle}{\|\mathbf{x}_2 - \mathbf{x}_1\|_{\mathcal{H}}} = 0. \quad (48)$$

In Euclidean spaces the Fréchet derivative is obviously the differential of  $f(\mathbf{x})$  at  $\mathbf{x}_1$ , in which case  $\nabla(\mathbf{x}_1)$  is the gradient and  $\langle \cdot, \cdot \rangle$  the Euclidean inner product. With a slight abuse of terminology, we will therefore call  $\nabla(\mathbf{x}_1)$  ‘the gradient’ even in more general Hilbert space settings.

Having thus defined an update direction  $\nabla(\mathbf{x})$  in quite general spaces, we are now in a position to define an algorithmic strategy to optimise  $f(\mathbf{x})$ . We again use a version of our trusty Iterative Hard Thresholding algorithm, but replace the update direction with  $\nabla(\mathbf{x})$ . With this modification, the algorithm might also be called the Projected Landweber Algorithm [16], and is defined formally by the iteration

$$\mathbf{x}^{n+1} = P_{\mathcal{A}}(\mathbf{x}^n - (\mu/2)\nabla(\mathbf{x}^n)), \quad (49)$$

where  $\mathbf{x}^0 = \mathbf{0}$  and  $\mu$  is a step size parameter chosen to satisfy the condition in Theorem 5 below.

### 3.1 Theoretical Performance Bound

We now come to the second main result of this paper, which states that, if  $f(\mathbf{x})$  satisfy the Restricted Strong Convexity Property, then the Iterative Hard Thresholding algorithm can find a vector  $\mathbf{x}$  that is close to the true minimiser of  $f(\mathbf{x})$  among all  $\mathbf{x} \in \mathcal{A}$ . In particular, we have the following theorem.

**Theorem 5.** *Let  $\mathcal{A}$  be a union of subspaces. Given the optimisation problem  $f(\mathbf{x})$ , where  $f(\cdot)$  is a positive function that satisfies the Restricted Strict Convexity Property*

$$\alpha \leq \frac{f(\mathbf{x}_1) - f(\mathbf{x}_2) - \text{Re}\langle \nabla f(\mathbf{x}_2), (\mathbf{x}_1 - \mathbf{x}_2) \rangle}{\|\mathbf{x}_1 - \mathbf{x}_2\|^2} \leq \beta, \quad (50)$$

for all  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{H}$  for which  $\mathbf{x}_1 - \mathbf{x}_2 \in \mathcal{A} + \mathcal{A} + \mathcal{A}$ . Let  $\mathbf{x}_{opt} = \arg\min_{\mathbf{x} \in \mathcal{A}} f(\mathbf{x})$  and assume that  $\beta \leq \frac{1}{\mu} \leq \frac{4}{3}\alpha$ , then, after

$$n^* = 2 \frac{\ln \left( \delta \frac{f(\mathbf{x}_{opt})}{\|\mathbf{x}_{opt}\|} \right)}{\ln 4(1 - \mu\alpha)}, \quad (51)$$

iterations, the IHT algorithm calculates a solution  $\mathbf{x}^{n^*}$  satisfying

$$\|\mathbf{x}^{n^*} - \mathbf{x}\| \leq \left( 2\sqrt{\frac{\mu}{1-c}} + \delta \right) f(\mathbf{x}_{opt}) + \|\mathbf{x} - \mathbf{x}_{opt}\|. \quad (52)$$

In the traditional Compressed Sensing setting, this result is basically that derived in [7].

### 3.2 Proof of the Second Main Result

*Proof of Theorem 5.* The proof requires the orthogonal projection onto a subspace  $\Gamma$ . The subspace  $\Gamma$  is defined as follows. Let  $\Gamma$  be the sum of no more than three subspaces of  $\mathcal{A}$ , such that  $\mathbf{x}_{opt}, \mathbf{x}^n, \mathbf{x}^{n+1} \in \Gamma$ . Let  $P_\Gamma$  be the orthogonal projection onto the subspace  $\Gamma$ . We write  $\mathbf{a}_\Gamma^n = P_\Gamma \mathbf{a}^n$  and  $P_\Gamma \nabla f(\mathbf{x}^n) = \nabla_\Gamma f(\mathbf{x}^n)$ . Note that this ensures that  $P_\Gamma \mathbf{x}^n = \mathbf{x}^n$ ,  $P_\Gamma \mathbf{x}^{n+1} = \mathbf{x}^{n+1}$  and  $P_\Gamma \mathbf{x}_{opt} = \mathbf{x}_{opt}$ .

We note for later that with this notation

$$\begin{aligned} \text{Re}\langle \nabla_\Gamma f(\mathbf{x}^n), (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle &= \text{Re}\langle P_\Gamma \nabla f(\mathbf{x}^n), (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle \\ &= \text{Re}\langle \nabla f(\mathbf{x}^n), P_\Gamma (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle \\ &= \text{Re}\langle \nabla f(\mathbf{x}^n), (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle \end{aligned} \quad (53)$$

and

$$\begin{aligned} \|\nabla_\Gamma f(\mathbf{x}^n)\|^2 = \langle \nabla_\Gamma f(\mathbf{x}^n), \nabla_\Gamma f(\mathbf{x}^n) \rangle &= \langle P_\Gamma \nabla f(\mathbf{x}^n), P_\Gamma \nabla f(\mathbf{x}^n) \rangle \\ &= \langle \nabla f(\mathbf{x}^n), P_\Gamma^* P_\Gamma \nabla f(\mathbf{x}^n) \rangle \\ &= \langle \nabla f(\mathbf{x}^n), \nabla_\Gamma f(\mathbf{x}^n) \rangle. \end{aligned} \quad (54)$$

We also need the following lemma.

**Lemma 6.** *Under the assumptions of the theorem,*

$$\left\| \frac{\mu}{2} \nabla_\Gamma f(\mathbf{x}^n) \right\|^2 - \mu f(\mathbf{x}^n) \leq 0. \quad (55)$$

*Proof.* Using the *Restricted Strict Convexity Property* we have

$$\begin{aligned}
\left\| \frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n) \right\|^2 &= -\frac{\mu}{2} \text{Re} \langle \nabla(\mathbf{x}^n), -\frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n) \rangle \\
&\leq \frac{\mu}{2} \beta \left\| \frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n) \right\|^2 + \frac{\mu}{2} f(\mathbf{x}^n) - \frac{\mu}{2} f(\mathbf{x}^n - \frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n)) \\
&\leq \frac{\mu}{2} \beta \left\| \frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n) \right\|^2 + \frac{\mu}{2} f(\mathbf{x}^n).
\end{aligned} \tag{56}$$

Thus

$$(2 - \mu\beta) \left\| \frac{\mu}{2} \nabla_{\Gamma}(\mathbf{x}^n) \right\|^2 \leq \mu f(\mathbf{x}^n), \tag{57}$$

which is the desired result as  $\mu\beta \leq 1$  by assumption.  $\square$

To prove the theorem, we start by bounding the distance between the current estimate  $\mathbf{x}^{n+1}$  and the optimal estimate  $\mathbf{x}_{opt}$ . Let  $\mathbf{a}_{\Gamma}^n = \mathbf{x}_{\Gamma}^n - \mu/2 \nabla_{\Gamma}(\mathbf{x}^n)$ . Because  $\mathbf{x}^{n+1}$  is the closest element in  $\mathcal{A}$  to  $\mathbf{a}_{\Gamma}^n$ , we have

$$\begin{aligned}
\|\mathbf{x}^{n+1} - \mathbf{x}_{opt}\|^2 &\leq (\|\mathbf{x}^{n+1} - \mathbf{a}_{\Gamma}^n\| + \|\mathbf{a}_{\Gamma}^n - \mathbf{x}_{opt}\|)^2 \\
&\leq 4\|\mathbf{a}_{\Gamma}^n - \mathbf{x}_{opt}\|^2 \\
&= 4\|\mathbf{x}^n - (\mu/2) \nabla_{\Gamma}(\mathbf{x}^n) - \mathbf{x}_{opt}\|^2 \\
&= 4\|(\mu/2) \nabla_{\Gamma}(\mathbf{x}^n) + (\mathbf{x}_{opt} - \mathbf{x}^n)\|^2 \\
&= \mu^2 \|\nabla_{\Gamma}(\mathbf{x}^n)\|^2 + 4\|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + 4\mu \text{Re} \langle \nabla_{\Gamma}(\mathbf{x}^n), (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle \\
&= \mu^2 \|\nabla_{\Gamma}(\mathbf{x}^n)\|^2 + 4\|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + 4\mu \text{Re} \langle \nabla(\mathbf{x}^n), (\mathbf{x}_{opt} - \mathbf{x}^n) \rangle \\
&\leq 4\|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + \mu^2 \|\nabla_{\Gamma}(\mathbf{x}^n)\|^2 \\
&\quad + 4\mu[-\alpha \|\mathbf{x}^n - \mathbf{x}_{opt}\|^2 + f(\mathbf{x}_{opt}) - f(\mathbf{x}^n)] \\
&= 4(1 - \mu\alpha) \|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + 4\mu f(\mathbf{x}_{opt}) \\
&\quad + 4[\|(\mu/2) \nabla_{\Gamma}(\mathbf{x}^n)\|^2 - \mu f(\mathbf{x}^n)] \\
&\leq 4(1 - \mu\alpha) \|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + 4\mu f(\mathbf{x}_{opt}).
\end{aligned} \tag{58}$$

Here, the second to last inequality is the RSCP and the last inequality is due to lemma 6.

We have thus shown that

$$\|\mathbf{x}^{n+1} - \mathbf{x}_{opt}\|^2 \leq 4(1 - \mu\alpha) \|\mathbf{x}_{opt} - \mathbf{x}^n\|^2 + 4\mu f(\mathbf{x}_{opt}). \tag{59}$$

Thus, with  $c = 4(1 - \mu\alpha)$

$$\|\mathbf{x}^n - \mathbf{x}_{opt}\|^2 \leq c^n \|\mathbf{x}_{opt}\|^2 + \frac{4\mu}{1-c} f(\mathbf{x}_{opt}), \tag{60}$$

so that, if  $\frac{1}{\mu} < \frac{4}{3}\alpha$  we have  $c = 4(1 - \mu\alpha) < 1$ , so that  $c^n$  decreases with  $n$ . Taking the square root on both sides and noting that for positive  $a$  and  $b$ ,  $\sqrt{a^2 + b^2} \leq a + b$ ,

$$\|\mathbf{x}^n - \mathbf{x}_{opt}\| \leq c^{n/2} \|\mathbf{x}_{opt}\| + 2\sqrt{\frac{\mu}{1-c}} f(\mathbf{x}_{opt}). \tag{61}$$

The theorem then follows using the triangle inequality

$$\begin{aligned}\|\mathbf{x}^n - \mathbf{x}\| &\leq \|\mathbf{x}^n - \mathbf{x}_{opt}\| + \|\mathbf{x} - \mathbf{x}_{opt}\| \\ &\leq c^{n/2}\|\mathbf{x}_{opt}\| + 2\sqrt{\frac{\mu}{1-c}}f(\mathbf{x}_{opt}) \\ &\quad + \|\mathbf{x} - \mathbf{x}_{opt}\|.\end{aligned}\tag{62}$$

The iteration count is found by setting

$$c^{n/2}\|\mathbf{x}_{opt}\| \leq \delta(\mathbf{x}_{opt}).\tag{63}$$

so that after

$$n = 2\frac{\ln\left(\delta\frac{f(\mathbf{x}_{opt})}{\|\mathbf{x}_{opt}\|}\right)}{\ln c},\tag{64}$$

iterations

$$\|\mathbf{x}^n - \mathbf{x}\| \leq \left(2\sqrt{\frac{\mu}{1-c}} + \delta\right)f(\mathbf{x}_{opt}) + \|\mathbf{x} - \mathbf{x}_{opt}\|.\tag{65}$$

□

### 3.3 When and Where is this Theory Applicable?

Since we first derived the result here, it has been shown that properties such as the Restricted Strict Convexity Property do indeed hold for certain non-linear functions such as those encountered in certain logistic regression problems [13]. These recent findings thus further strengthen the case for a detailed study of non-convexly constrained non-linear problems and the derivation of novel methodologies for their solution.

It may thus seem tempting to use this theory also in a non-linear Compressed Sensing setting, where we would have  $f(\mathbf{x}) = \|\mathbf{y} - \Phi(\mathbf{x})\|_B^2$ , where  $\|\cdot\|_B$  is some Banach space norm and where  $\Phi(\cdot)$  is some non-linear function<sup>1</sup>. If this  $f(\mathbf{x})$  would satisfy the Restricted Strict Convexity property, then the Theory in the second part of this paper would indeed tell us how to solve the non-linear Compressed Sensing problem.

Unfortunately, it is far from clear yet under which conditions on  $f(\mathbf{x}) = \|\mathbf{y} - \Phi(\mathbf{x})\|_B^2$  Restricted Strict Convexity type properties hold. Indeed, the following lemma shows that such a condition cannot be fulfilled in general for Hilbert spaces.

**Lemma 7.** *Assume  $\mathcal{B}$  is a Hilbert space and assume  $f(\mathbf{x})$  is convex on  $\mathcal{A} + \mathcal{A}$  for all  $\mathbf{y}$  (i.e. it Satisfies the Restricted Strict Convexity Property), then  $\Phi$  is affine on all subspaces of  $\mathcal{A} + \mathcal{A}$ .*

---

<sup>1</sup>This was indeed the setting proposed in [11].



*Proof.* The proof was suggested by an anonymous reviewer of the earlier version of this manuscript [11] and uses contradiction. Assume  $\Phi$  is not affine on any subspace of  $\mathcal{A} + \mathcal{A}$ . Thus, there is a subspace  $\mathcal{S} = \mathcal{A}_i + \mathcal{A}_j$ , and  $\mathbf{x}_n \in \mathcal{S}$ , such that for  $\mathbf{x} = \sum_n \lambda_n \mathbf{x}_n$ , where  $\sum_n \lambda_n = 1$  and  $0 \leq \lambda_n$ , we have  $\sum_n \Phi(\mathbf{x}_n) - \Phi(\mathbf{x}) \neq \mathbf{0}$ . Now by assumption of strong convexity on  $\mathcal{S}$ , we have (using  $\mathbf{y}_n = \Phi(\mathbf{x}_n)$  and  $-\bar{\mathbf{y}} = \mathbf{x}$ )

$$\begin{aligned} 0 &\leq \sum_n \lambda_n \|\mathbf{y} - \Phi(\mathbf{x}_n)\|^2 - \|\mathbf{y} - \Phi(\mathbf{x})\|^2 = \sum_n \lambda_n \|\mathbf{y} - \mathbf{y}_n\|^2 - \|\mathbf{y} - \bar{\mathbf{y}}\|^2 \\ &= 2\langle \mathbf{y}, \bar{\mathbf{y}} - \sum_n \lambda_n \mathbf{y}_n \rangle + \sum_n \lambda_n \|\mathbf{y}_n\|^2 - \|\bar{\mathbf{y}}\|^2. \end{aligned} \quad (66)$$

where the inequality is due to the assumption of convexity. But the above inequality cannot hold for all  $\mathbf{y}$  (it fails for example for a multiple of  $-(\bar{\mathbf{y}} - \sum_n \lambda_n \mathbf{y}_n)$ ). Thus  $\Phi$  needs to be affine on the linear subsets of  $\mathcal{A} + \mathcal{A}$ .  $\square$

Whilst this implies that the property cannot hold in Hilbert spaces for non-affine  $\Phi$  and *all*  $\mathbf{y}$ , it does not preclude the possibility that it could hold for specific observations  $\mathbf{y}$ . This would not allow us to build a general signal recovery framework, but might still allow us the recovery of a subset of signals. Thus, for the non-linear Compressed Sensing problem in Hilbert space, the Restricted Isometry Property of the Jacobian of  $\Phi(\mathbf{x})$  together with the ability to construct a good linear approximation of  $\Phi(\mathbf{x})$  seem to be the more suitable tools to study recovery performance. Nevertheless, for certain other non-convexly constrained non-linear optimisation problems, such as those addressed in [13], the Restricted Strict Convexity Property might be the more appropriate framework. Whilst there are many similarities between these requirements and they both boil down to the same RIP property in the linear setting, it remains to be seen what the exact relationship is between these two measures in general non-linear problems.

## 4 Conclusions

Compressed Sensing ideas can be developed in much more general settings than considered traditionally. We have shown previously [14] that sparsity is not the only structure that allows signals to be recovered and that the finite dimensional setting can be replaced with a much more general Hilbert space framework. In this paper we have made a further important generalisation and have introduced the concept of non-linear measurements into Compressed Sensing theory. Under certain conditions, such as the requirement that the Jacobian of the measurement system satisfies a Restricted Isometry Property, then the Iterative Hard Thresholding algorithm can be used to recover signals from a non-convex constraint set with similar error bounds to those derived in Compressed Sensing.

In the second part of this paper we have then looked at the related and in some sense more general setting of non-linear optimisation under non-convex constraints. Here we have looked the Restricted Strict Convexity Property as a tool to study recovery performance and it was shown that that this condition is indeed sufficient for the Iterative Hard Thresholding to find points that are near the optimal solution.

## Acknowledgment

This work was supported in part by the UKs Engineering and Physical Science Research Council grants EP/J005444/1 and D000246/1 and a Research Fellowship from the School of Mathematics at the University of Southampton.

## References

- [1] E. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, 2006.
- [2] D.L. Donoho, “For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution,” *Communications on Pure and Applied Mathematics*, vol. 59(6) pp. 797–829, 2006
- [3] E. Candès and M. B. Wakin “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25(2), pp. 21–30, 2008
- [4] E. Candès, “The restricted isometry property and its implications for compressed sensing,” *Compte Rendus de l’Academie des Sciences*, Paris, Serie I, 346 589–592, 2008.
- [5] D. Needell and J. Tropp, “COSAMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comp. Harmonic Anal*, vol. 26, pp. 301–321, 2008.
- [6] W. Dai and O. Milenkovic, “Subspace pursuit for compressed sensing: Closing the gap between performance and complexity,” *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.
- [7] T. Blumensath and M.E. Davies “Iterative Hard Thresholding for Compressed Sensing,” *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009
- [8] M. Mishali, Y.C. Eldar “Blind Multi-Band Signal Reconstruction: Compressed Sensing for Analog Signals,” *IEEE Transactions on Signal Processing*, vol. 57(3), pp. 993–1009, 2009

- [9] D. Goldfarb, S. Ma “Convergence of fixed point continuation algorithms for matrix rank minimization,” *arXiv:09063499v3*, 2010.
- [10] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Transactions on Information Theory*, vol 56(4), pp. 1982–2001, 2010
- [11] T. Blumensath “Compressed Sensing with non-linear observations,” <http://eprints.soton.ac.uk/164753/>, Oct. 2010
- [12] W. Xu, M. Wang, A. Tang “Sparse Recovery from Nonlinear Measurements with Applications in Bad Data Detection for Power Networks,” *arXiv:1112.6234v1*, 2011
- [13] Bahmani S., Raj B., Boufounos P. “Greedy Sparsity-Constrained Optimization,” *arXiv:1203.5483v1*, 2012
- [14] T. Blumensath, “Sampling and reconstructing signals from a union of subspaces,” *IEEE TRans. on Information Theory*, vol. 57(7), pp. 4660–4671, 2011
- [15] A. Kyrillidis and V. Cevher ”Combinatorial Selection and Least Absolute Shrinkage via the CLASH Algorithm” *preprint*, 2011
- [16] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, 2000.
- [17] S. Negahban, P. Ravikumar, M. J. Wainwright and B. Yu. “A unified framework for the analysis of regularized  $M$ -estimators’,’ in *Advances in Neural Information Processing Systems*, December, 2009. Vancouver. Canada.